![Global Alliance for Responsible Media](GARM logo)

GARM: 3 Years of Progress

Uncommon Collaboration and its Impact on Brand Safety

A look back and reflecting on our journey

Three years ago, the WFA led 16 organizations in the launch of the Global Alliance for Responsible Media (GARM) at the Cannes Lions Festival of Creativity. GARM is a cross-industry initiative, launched by brands, to remove harmful content from ad-supported digital and social media. That challenge is big, however GARM has been able to develop a unique structure with focused areas for collaboration that have driven meaningful impact.

GARM's launch was propelled forward by *uncommon collaboration,* a unique way of working recognizing that all sectors of the advertising industry and companies benefit from partnering to create new brand safety standards and solutions that could be accepted industry-wide, where there had been no established protocols.

From an initial set of 16 launch companies, we've grown to 122 members (61 advertisers, six agency holding companies, 11 media platforms, nine ad tech companies, and 35 industry associations). GARM has six active

working groups, powered by 80 media leaders from a larger GARM Community. The Community meets monthly to review the progress of Working Groups and share best practices and thought leadership.

Along with our launch focus of advertisers, agencies, platforms and industry associations, we've set focused areas for engagement:

- Bringing advertising technology companies into a Solutions Developers Working Group to help us drive consistency in implementation and faster speed to market;

- Broadening membership to National Advertiser Associations to help us embed GARM work at the local level;

- Formalizing our ways of working with NGOs via a consult group to ensure our work isn't insular to the advertising industry.

A view on digital media safety: our right to play & win

Through the work of the GARM Steer Team, we've brought into focus how digital media safety should be viewed holistically. The GARM Steer Team's view on digital media safety is framed by two broad questions:

Is the platform safe for consumers?	Does the service cause harm? Is engagement positive? Is support universal? Is moderation fair? Is there choice and control?	Place human safety at the core to eliminate harm from the attention flywheel	PLATFORM SAFETY
Is the platform suitable for advertising?	Is content categorized? Are there placement controls? Is there transparency in where ads appeared?	Make digital media safety practices analogous to other media	BRAND SAFETY

The first area of consideration is **Platform Safety,** which focuses on the product, essential technology design and oversight. These are individual media platform responsibilities that sit well outside advertiser control and competency. These also raise questions that are fundamentally local and regulatory in nature. Areas outside of advertiser control are content legality, freedom of expression and algorithmic oversight. These are areas where we need regulators to step in and define the right balance between consumer protection and freedom of expression. Further, advertisers don't have full competency in areas like technology design and algorithms where we need technical regulation as seen in markets like the EU via the Digital Services and Digital Markets Acts and in Australia with the eSafety Commissioner.

Advertisers have had a clear stake in **Brand Safety,** and with GARM we are able to overcome fragmentation within the advertiser, agency and platform communities to set a direction forward. This has enabled us to effectively establish safety standards for where ads show up. Advertiser standards on content safety work best for our industry applied at a global level to drive impact, and GARM's global scope of influence positions us to continue leading here.

GARM's focus remains on Brand Safety, which is centered on monetization safety – how advertising investment is steered away from harmful content and behaviors. Since our start, we've been able to drive positive, forward momentum around four core areas we set out in the GARM Charter, launched at the WEF Davos Summit in January 2020. They are:

- Common Definitions
- Common Metrics
- Common Tools
- Independent Verification
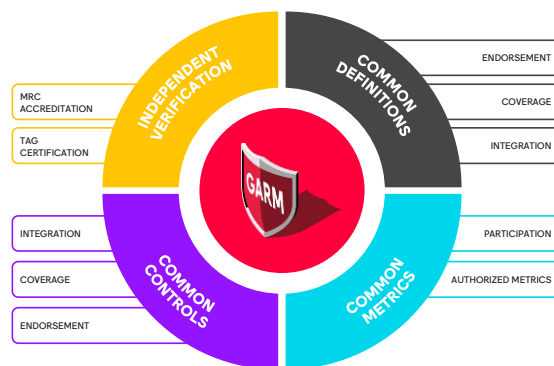
With that understanding of where we have direct impact on Brand Safety, and indirect influence on Platform Safety – our hypothesis is that transparency, control and accountability will better allow the advertising industry to reward positive content and engagement.

All our solutions are meant to improve upon that –brands should be able to invest in content that aligns with their values and purpose, and with proper tools and partnerships in place there should be no surprises.

Common Definitions:
Setting the Limits for Advertising Support

What we've delivered so far – The GARM Brand Safety Floor + Suitability Framework

Eliminating harmful content from advertising campaigns requires a shared understanding of what are sensitive topics, and what are the limits for advertising support. Prior to our work in this area, platforms, advertisers and agencies had separate views and vocabulary on harmful content. In September 2020, GARM started with the 4A's APB initial proposals and enhanced them through a multistakeholder process and in collaboration with GARM's NGO Consult Group. From there, GARM drove an agreement across advertisers, agencies and platforms to set a framework that limits advertising support for harmful content, through the Brand Safety Floor, while providing for a Suitability Framework to manage advertising placement in sensitive content categories, while acknowledging critical nuances across platforms and formats.

What's new today – GARM-inspired ad tech solutions and a new category for Misinformation

Two years from the launch of the GARM Brand Safety Floor + Suitability Framework, we are proud to show that this solution is impacting how brands set strategies, how media agencies build media buys, and how platforms and ad tech partners structure their tools. Because of this work, advertisers have more tangible control and transparency over monetized content. We are also proud to announce that we have added Misinformation into this framework, in coordination with our work with the European Commission and several of our NGO partners. This important

work helps to solidify individual GARM member work and multistakeholder collaboration with regulators. The new standard is designed to provide a structure for demonetizing harmful misinformation and to build on the success that the framework has already delivered.

What's next tomorrow – Democratizing application in local markets

As noted, much of our work is global in nature. However, we recognize that this framework in the hands of more advertising buyers and sellers will fundamentally further the demonetization of harmful content through simple demand and supply marketplace dynamics. To that end, we are working through WFA's National Associations Council to help translate, calibrate and embed it via local educational efforts with national advertiser associations.

GARM: Brand Safety Floor + Suitability Framework

Common Metrics: Tracking Industry Efforts

What we've delivered so far – The GARM Aggregated Measurement Report

Following our agreements on Definitions, our focus turned to driving transparency through tracking industry progress in removing harmful content from advertising. In April 2021, GARM developed a common framework to assess industry progress in removing harmful content from advertising-supported media. In partnership with seven member platforms we defined four core questions and eight authorized metrics to drive transparency for the advertising industry. As a result of the GARM Aggregated Measurement Report, new metrics have been shared that have never been available before ranging from YouTube's Violative View Rate and Advertising Safety Error Rate to Meta's reporting on Prevalence of Hate Speech to Snap's Violative View Rate to Pinterest's reporting on removal of Misinformation content by views and to Twitter's sharing of Violative Impressions Rate.

What's new today – Increased participation and first accreditations

Since the launch of this effort, we've had more platforms join, and the measurement best practices and authorized metrics are shaping new work from new platforms. We've also been consistent in our call to have independent verification of metrics for monetization and transparency reporting. We're pleased to share that platforms are responding to our calls to have their monetization safety metrics audited. YouTube is the only platform at present to have their monetization safety metrics accredited by the MRC – this is a significant independent verification of the platform's safety for advertising and should help improve the confidence in the safety of YouTube's monetized content. We urge other platforms to follow. We also note that Meta have completed a first-party audit via EY of their transparency reporting that helps attest to the accuracy of their own processes.

What's next tomorrow – Increased disclosure for local markets

What's next in this area is taking the global metrics we have and gaining regional and language-level insights. We must help the industry go beyond global understandings to local trust-building transparency. Leading platforms are already exploring how to provide more specific metrics and an overview of global sampling methodologies, and we anticipate being able to report out a roadmap with Volume 5 of the report in less than a year from now. We also want to leverage work already done by platforms with regulators and NGOs to help understand safety incidents with a pertinent lens for advertisers: how many people were reached by the harmful content and was it supported by advertising?

Question	Authorized Metric	YouTube	Facebook	Instagram	Twitter	TikTok	Pinterest	Snapchat	Twitch
How safe is the platform for consumers?	Prevalence Violative View Rate	Authorized Metric	Authorized Metric	Authorized Metric	Next Best Measure	Next Best Measure	Next Best Measure	Authorized Metrics	Next Best Measure
How safe is the platform for advertisers?	Advertiser Safety Error Rate or Prevalence	Authorized Metric	Authorized Metric	Authorized Metric	Next Best Measure	Next Best Measure	Next Best Measure	Authorized Metric	Authorized Metric
How effective is the platform at enforcing its safety policies?	Removals of violating content	Authorized Metric	Authorized Metric	Authorized Metric	Authorized Metric	Next Best Measure	Authorized Metric	Authorized Metric	Authorized Metric
	Removal of violating accounts by views	Authorized Metric	Authorized Metric	Not Submitted	Next Best Measure	Authorized Metric	Authorized Metric	Authorized Metric	Authorized Metric
	Removal of violating accounts	Authorized Metric	Authorized Metric	Not Submitted	Authorized Metric	Not Submitted	Authorized Metric	Authorized Metric	Authorized Metric
How responsive is the platform in correcting mistakes?	Appeals (pieces of content)	Authorized Metric	Authorized Metric	Authorized Metric	Not Submitted	Not Submitted	Authorized Metric	Not Submitted	Authorized Metric
	Reinstatements (pieces of content)	Authorized Metric	Authorized Metric	Authorized Metric	Not Submitted	Not Submitted	Authorized Metric	Not Submitted	Authorized Metric

GARM Global Alliance for Responsible Media

Common Tools: Driving widescale safety for ad placement

What's new today — The GARM Adjacency Standards Framework

In January 2021 we started work on the GARM Adjacency Standards Framework via a dedicated Working Group. Designed to serve as a companion to the GARM Suitability Framework, these standards provide advertising industry participants with a common methodology for evaluating the brand suitability of an ad placement relative to an ad's position to nearby content (i.e. "adjacency") within specific media formats. The Adjacency Standards Framework works within the confines of the GARM Suitability Framework – where sensitive content can be supported by advertising, and effectively allows advertising buyers and advertising sellers more consistency and control over placements relative to sensitive content.

This Working Group assessed a series of existing tools and consumer research, and also conducted research within the GARM Community to assess needs across advertisers and agencies. We're happy to share that we have now defined adjacency standards for News Feed, Stories, In-stream Video, In-stream Audio, and Display overlay. These are minimum standards and we are eager to see the industry take these up as a means of managing brand suitability. Given the recent incidents involving Livestream formats, we are working internally to ensure that there is a robust Safety Floor in place, before advancing those formats into a Suitability Framework.

What's next — Moving from Adjacency Standards and building toward Controls

With the standards for many formats defined, and livestream as a fast follower, this GARM Working Group will then pivot to working together to educate the marketplace on current and planned development. This Working Group will be collaborating with platforms and ad tech partners take these standards and putting them into practice. Once they are available, GARM's goal is to have them assessed as part of existing independent audits.

FORMAT		ADJACENCY STANDARD	MINIMUM ADJACENCY UNIT	NOTES
Feed	This covers content (text, video, image, audio) that is featured in a newsfeed or timeline environment irrespective of the screen being mobile or computer.	Spatial	+/- 1	• Adjacency controls should apply to individual and group feed and timeline (e.g., lists and groups that are public and/or private) • Comments on videos are not a focus for suitability controls – but platform must be able to uphold Floor on videos
Stories	This covers sequenced content from a single creator in a carousel environment, where ads may appear within or between such segments.	Spatial	+/- 1	• Comments on stories are not a focus for suitability controls – but platform must be able to uphold Floor on videos
Video: In-stream	This is prerecorded video content that is uploaded to a website or platform that features ads before, in between or after specified video content	Temporal	+/- 0 Directly Adjacent	• Adjacency solutions and controls should apply to Pre/Mid/Post and Parallel ad units • Comments on videos are not a focus for suitability controls – but platform must be able to uphold Floor on videos
Audio: In-stream	This is pre-recorded audio content that is uploaded to a website or platform that features ads before, in between or after specified video content.	Temporal	+/- n Same as ad unit length (n = ad length)	• Adjacency controls should apply to Pre/Mid/Post and Parallel ad units • Comments on content are not a focus for suitability controls – but platform must be able to uphold Floor on audio content

Independent Verification: Building trust in process and operations

What's new today – More platforms with more accreditations

Digital platforms have become the cornerstones of advertising. We must have trust and transparency at the core of the business, especially in safety. We've aligned two audit standards – TAG Brand Safety Certification which looks at process, and MRC Content Level Brand Safety Accreditations which looks at definitions, implementation, and reporting. We're pleased to share that all GARM platforms are TAG-certified. We are happy to see that YouTube has led the way in being the only platform at present to have earned MRC's Content Level Brand Safety Controls certification – inclusive of operations and reporting on monetization. We are encouraged by public commitments and continued steps made by other platforms like Meta and Twitter.

What's next – Getting Local on Independent Verification Results

We know that brand safety is rooted in culture, which is why it is important to take the results of global audits and understand them at a local and language level. Again, this is something we will do with WFA's National Associations Council. This will allow us to better line up advertising growth at a local market-level and safety for users. We believe that as markets develop for advertising revenue, there should be commensurate investments into safety.

GARM's go-forward

Demonetizing harmful content online is a big challenge, and content continually shifts as the culture evolves. We are starting to see impact in our work, and we do believe that GARM is an effective forum to address brand safety. And we need to acknowledge two provocations:

Provocation 1: Brand Safety isn't a substitute for Platform Safety

Platforms have become a mainstay in consumers' lives and in the industry. However, platforms have increasingly been forced to make hard decisions on freedom of expression, consumer safety and technology transparency. We are supportive of regulation that works both in the interest of society and in the interest of industry to set respective floors and the right duty of care that all users deserve, creating robust and consistent thresholds for platform safety. We have been, and continue to be, supportive of progressive, comprehensive regulation that creates a duty of care around design, resourcing, oversight, and moderation. Freedom of expression and safety are not mutually exclusive. Similarly, while we recognize that regulators want to avoid stifling development of new platforms for safety requirements, every user's safety should be protected. We believe this helps platforms have common ground on safety requirements, recognizing they operate in a competitive arena.

Provocation 2: We must get the future right by being proactive

GARM was created in reaction to a lack of effective and holistic safeguards that inadvertently had advertising funding harmful content. With GARM's work underway and impact being seen, we must now help the industry understand safety requirements before commercialization begins in the metaverse. We're being asked by our members to start on this journey, and we're here in Cannes to recommit to our mission and renew it and scale it to new spaces. We must ensure that advertising is aligned with sustainable and responsible growth models. We anticipate sharing a plan for this in conjunction with our work with regulators, our NGO Consult Group, and partners like the World Economic Forum. We are challenging ourselves to share progress towards a framework in January 2023.

While we are proud to share our progress three years on, we are clear; our impact shouldn't distract us from the challenge and opportunities at hand. It is our intent to build upon our accomplishments to date and this truly uncommon coalition to scale our efforts to engage our partners locally and define proactive practices that will future proof our industry for the good of the digital media ecosystem and our society.

To our members and our supporters – we are humbled by your commitment and contributions to our work.

To those interested in joining GARM, we invite you to join a community of the committed and like-minded.

Contact information

Rob Rakowitz, Co-founder & Initiative Lead - Global Alliance for Responsible Media
r.rakowitz@wfanet.org

About the World Federation of Advertisers

The World Federation of Advertisers (WFA) is the voice of marketers worldwide, representing 90% of global marketing communications spend – roughly US$900 billion per annum – through a unique, global network of the world's biggest brand owners and national advertiser associations in more than 60 markets. WFA champions more effective and sustainable marketing communications. More information at wfanet.org.

About the Global Alliance for Responsible Media

The Global Alliance for Responsible Media (GARM) was formed to identify specific collaborative actions, processes and protocols for protecting consumers and brands from safety issues. Alliance members will work collaboratively to identify actions that will better protect consumers online, working towards a media environment where hate speech, bullying and disinformation is challenged, where personal data is protected, and used responsibly when given, and where everyone is, especially children, better protected online. Alliance members acknowledge their collective power to significantly improve the health of the media ecosystem. More information at wfanet.org/GARM.